DOCUMENT RESUME

ED 307 696 EA 021 020

AUTHOR Folger, John
TITLE Three Policy Relevant Findings from Project STAR.
PUB DATE Mar 89
NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989). For related document, see EA 021 021.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Class Size; *Cost Effectiveness; *High Risk Students; Primary Education; *Small Classes; *Socioeconomic Influences; Standardized Tests
IDENTIFIERS *Project STAR

ABSTRACT

This paper addresses three questions policymakers may have concerning the results of the Project STAR data. First, when reducing class size, is there a "turning point" at which class size reductions lead to increased gains in student achievement? Second, do small classes help low socioeconomic status "at risk" students more than they help high socioeconomic status students? Third, what is a useful way of determining whether the gains from reducing class size are worth the costs involved? Appended are four tables summarizing standard achievement test scores. (SI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED307696

THREE POLICY RELEVANT FINDINGS FROM PROJECT STAR

John Folger

Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, March 27-31, 1989.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Folger

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 021 020

BEST COPY AVAILABLE

THREE POLICY RELEVANT FINDINGS FROM PROJECT STAR

John Folger

The primary results of the analysis of the Project STAR data have been summarized in the paper by Dr. Finn. This paper looks at three questions which policy makers are interested in:

First, when reducing class size, is there a "turning point" at which class size reductions lead to increased gains in student achievement?

Second, do small classes help low socioeconomic (SES) "at risk" students more than they help high SES students?

Third, what is a useful way of determining whether the gains from reducing class size are worth the costs involved?

Is there a turning point at which class size reductions lead to increased gains in student achievement?

Glass, in his meta analysis of class size studies, identified a class size of 15 as the point below which achievement gains accelerated for each student reduction in class size. His trend line had a noticeable bend at about 15 students. It was on the basis of this study that the legislation for Project STAR specified that small classes should have an average of 15 students. It might have been useful if the drafters of the legislation could have read Robinson (1985) and Walberg's (1988) criticism of the Glass analysis. Glass included tutorials (1 on 1) as a class size, and that alone could have produced the curvilinear relationship that Glass reported.

The question of whether the relationship between class size and achievement is linear, or whether it is curvilinear was not answered satisfactorily by Glass's meta analysis. Project STAR had the data necessary

ED 321 020

to answer the question, at least for classes in the size range 13 to 26.

There were 328 kindergarten, 346 first grade and 340 second grade classes, distributed in size across all the sizes from 13 to 26. For example, in second grade there were 14 classes with 13 students, 17 with 14, 32 with 15, etc. The distribution by size was bimodal, with one peak at 15 and another at 22.

Table 1 shows correlations between reading, math, and listening achievement scores and individual class sizes for each grade. They are all low and in the expected negative direction, that is as class size decreases, achievement scores increase. These correlations indicate that class size "explains" about 3 to 5 percent of the variance in average class achievement.

We tested all of the correlations to see if they deviated from linearity, and none of them did. What this means is that reducing class size from 25 to 20 will have about the same effect as reducing it from 22 to 17, or from 20 to 15. There is no "turning point" where decreasing class size by one student has a bigger effect than at any other point.

Correlations were computed for three subscores in the Stanford Achievement Test for kindergarten, first and second grade. For the achievement tests that were used, and for K, 1 and 2, class sizes between 13 and 26, the relation between class size and achievement is linear. Policy makers can expect about as much achievement gain for a reduction from 25 to 20, as for a reduction from 20 to 15.

Glass did not offer any conceptual basis for expecting the relation of achievement to class size to be nonlinear, and an examination of the literature did not uncover any hypothesis that would explain a sudden increase in achievement at any particular class size. Most of the explanations for

class size effect suggest that the effect would be inversely proportional to the number of children in the class, and while this is not a linear function, over the range of 13 to 26, it isn't far from a straight line.

Achievement by Socioeconomic Status

Students from low socioeconomic backgrounds are generally more "at risk" and public policy has been focused on ways to give these students additional help (resources) to enable them to be as successful as high SES students.

One approach to enhancing low socioeconomic students' performance is to put them in a smaller class, or a class with an aide, to provide them with more teacher attention. Project STAR provides an opportunity to examine the effect of lowering class size on students from different SES levels.

The measure of socioeconomic status that was available was free or reduced cost lunch (low SES), or pay full price (high SES). Eligibility for free lunch is determined by parental income in relation to family size. In second grade there were 3,044 high SES students, who had average reading scale scores of 600, while the 3,051 low SES students had average scores of 567, 33 points lower. This difference is about three times as large as the score difference between small and regular classes in second grade. It also represents a grade equivalent difference of about four months. It is clear that if low SES students could get a better start and close some of the gap with high SES students, the overall achievement level would be raised, and later problems of dropouts and low performance might be minimized.

Low SES students did better in reading and math in small classes than in regular classes in each type of school, which is consistent with findings reported earlier that small classes produce bigger gains than regular classes. (See Table 2.) An important question is not whether small classes help low

SES students, but whether they help them more than they help high SES students. For each school type the score advantage of small classes over regular classes was determined separately for the low SES and high SES students. If the "advantage" was larger for low SES students, then a small class helped them more. (See Table 2 for reading and Table 3 for math.) In two of the four school types (inner city and rural) the small class helped the reading scores of the high SES students more than it helped the low SES students. In urban schools there was no difference (less than 1.0 points) and in suburban schools small classes aided the low SES students more.

The same pattern existed for math scores. Low SES students were helped more than high SES students by small classes in suburban schools, but they were helped less than high SES students in the other three school types.

Another way to look at the effects of a small class on low SES students is to see whether the gains from the end of kindergarten to the end of second grade in the small classes have been greater for low or high SES students. In every school type, and for both math and reading, gains have been greater in small class for the high SES students, as compared with low SES students. High SES students were ahead of low SES students at the end of kindergarten, and they are even further ahead by the end of second grade. (See Table 4.)

Another way of comparing the effect of small classes is to see if the gains from kindergarten to second grade of low SES students in small classes are as great or greater than the achievement gains of high SES students in regular classes. If that is the case, policy makers could reduce class size in schools with high proportions of low SES students to provide the students with extra resources of teacher time which could help their achievement level. In reading in two of the four school types (inner city and urban) achievement

gains of low SES students in small classes were almost as great as high SES students in regular classes. In the other two school types, suburban and rural, high SES students in regular classes had larger achievement gains than low SES students in small classes. (See Table 4.) In math, in suburban and urban schools, low SES students in small classes had slightly greater gains than high SES students in regular classes; in inner-city and rural schools, the high SES students in regular classes had larger gains than the low SES students in small classes. (See Table 4.)

To summarize all of these socioeconomic comparisons, there is no consistent evidence from the analysis of achievement test scores to indicate that small classes help low SES students more than they help high SES students. In a majority of the comparisons both of level and of gain, the high SES students are helped more by small classes than are low SES students.

One problem with the analysis of gain scores just presented is that the gain scores are computed from class averages and not from individual scores. The movement of students in and out of the classes could have affected the class average and the gain score. The next step in the analysis of SES differences will be to calculate individual gain scores K to 2 for those students who have been in the project in the same class type (small, regular, or regular/aide) for all three years, and then average the individual gain scores for school and class type and SES. If this produces the same general findings that high SES students are helped as much or more than low SES students by small classes, one possible explanation for these differences will have been removed.

One possible explanation for the finding that small classes help high SES students as much as low SES students is that teachers don't concentrate

the extra time per student they have in a small class on the low SES students, but instead, spread it over all students in the class. If the objective is to concentrate more attention on the low achievers, teachers will have to alter their teaching strategies to accomplish this. The strategies that will work in a class with 3 or 4 low SES children may need to be different than those in a class with 75-80 percent of low SES children. To explore some of these possibilities we plan additional analyses.

Individual school level data can be analyzed to identify individual teachers who (based on scores and gain scores) are effective in closing the achievement gap between low and high SES students. They can then be interviewed and observation data (if available) and other process data can be examined to see if we can determine how they are achieving results. Do they have a specific goal of helping low SES students, or is the result simply a byproduct of a very effective teacher?

What is the Cost Effectiveness of Reduced Class Size

Estimation of cost effectiveness of a treatment such as reducing class size requires a measure of cost, a measure of effectiveness, and a standard against which to compare both cost and effectiveness. For Project STAR, the standard of effectiveness we chose was the average performance of students in the state on achievement measures in reading and math.

The cost standard was the average monthly cost of instruction in schools across the state. A more precise measure would have been grade-level costs, but estimates of these are not available. Since our cost estimates are more precise than our estimates of effectiveness, greater precision in cost estimation is less important.

The estimate of effectiveness was achievement scores of students. For

small classes, effectiveness was the average "advantage" of student scores in small classes compared with regular classes. This advantage could be expressed in months of advantage. For example, if small class students were on average a month ahead of regular class students in reading in kindergarten, the advantage of a small class was one month. Since it took a year to produce that advantage, the extra cost of teaching the small class for a year was the cost of producing that advantage. Achievement in reading and math is an important, but not the only important objective of schooling. In the first grade teacher logs report an average of about 2-1/4 hours spent on reading and language arts, and an average of about 53 minutes on math. These subjects take up between 1/2 and 2/3 of the total school day. For purposes of cost allocation, we make the assumption that there was about the same use of time in small classes and regular classes, and in fact the allocation of time reported on the logs was quite similar.

The small class cost was then compared with the average cost of providing the same amount of growth in achievement in a regular class that the small class advantage represents. For example, if the small class is a month ahead in reading, the comparison cost would be the average cost of producing a month's gain in a regular class. In Tennessee students in kindergarten have achievement scores in both reading and math that are slightly above national norms, and they are still above national norms at the end of second grade. So we made the assumption that on the average, students in the primary grades made a month's achievement gain for each month they are in school.

If we allocate half the cost of one month's instruction to the "production" of one month's gain in reading and a little more than one month's gain in math, we will have an average cost in a normal-sized class of

producing this much gain. We can then compare the cost of reducing class size or having a teacher aide (or any other "improvement" in the program) with the usual average cost of producing this much learning in the average classroom in Tennessee.

The estimate that half the cost of an average month should be allocated to the reading and math achievement objectives is arbitrary, but it is consistent with the average amount of the classroom day that is devoted to these objectives. As an alternative, the entire cost could be allocated to these two objectives (math and reading) on the logic that the other use of time in the school day is really designed to support these "primary" objectives. For the purpose of comparing small and regular classes' costs, it doesn't matter which cost allocation is used, just so the same assumptions are made for both comparisons.

The 1985-86 per-pupil expenditure average in Tennessee is $2,311. (This is an average over all grades; an average for each individual grade is not available.) An average month's cost per pupil (nine month's school year) is $257. This is an estimate of the average cost of producing an average gain of about 0.1 grade equivalent in achievement in reading and math, when all costs are allocated to these two objectives. If only half the costs are allocated to these objectives, the cost of the gain is $128.

To estimate the cost of producing this effect size with a small class, we have to estimate how much more, per pupil, a small class costs. Slightly more than half the total per-pupil cost of $2,311 is represented by the teacher's salary. The average salary plus benefits per teacher in Tennessee was $26,962 in 1985-86. There was a 22.5 to 1 teacher pupil ratio in kindergarten in 1985-86; the teacher salary plus benefits cost per pupil is

$1,198. We assume that this is the only significant cost difference per pupil associated with class size. Transportation costs, administrative costs, supplies and equipment would all be the same. (This underestimates the costs of a small class because the cost of building maintenance would increase because more classrooms would be required.) In this analysis, capital costs were ignored and only the teacher salary cost difference was added.

If class size is reduced to 15 to 1, with the same average teacher salary plus benefits, teacher salary costs would be $1,797 and total per-pupil costs would be $2,910, or $599 more than the average cost in 1985-86. If half the cost is allocated to reading and math instruction, the increased cost per pupil would be $299.50.

The average current cost per pupil of producing a month's gain in reading and math in Tennessee	$128
The average cost of producing the extra month's gain in reading and math in the small class	$299

At the end of the first grade, pupils in the small class were approximately two months ahead of pupils in the regular classes in both reading and math. One month of the difference occurred in kindergarten, and should be charged to kindergarten costs, and the other month occurred in first grade, and is charged to first grade costs.

This produces the same cost effectiveness picture in first grade as in kindergarten:

Average cost per pupil of producing a month's gain in first grade	$128
Average cost of an extra month's gain in small class in reading and math in the first grade	$299

Discussion

All of these estimates should be considered as tentative; they are based on rough estimates of both "effectiveness" and cost. They ignore other possible benefits of a small class. For example, improved socialization of children to the school environment and the possibility that the advantage of small classes in other subject areas (if we had tested them) would have been even greater than in math or reading.

The analysis also assumes that each month's achievement in reading and math is of equal value, while we might assume that gains in achievement are of increasing value. For example, a month of additional achievement above the average learning in a month could be more valuable than additional achievement below the average. However, we might believe that it was more important to bring low achievers up to average, which might lead to a different valuation.

Another potential benefit that is not included in this analysis is the value of teachers' job satisfaction. A more satisfied teacher may stay in teaching longer and reduce the costs of bringing new personnel into the system. To the extent that teachers' morale and job satisfaction is increased with a small class, some dollar value might be assigned to this to reduce the differential cost of producing more student learning.

If the estimates we have developed are of the right approximate magnitude, it indicates that achievement differences of 2 to 2-1/2 months between small and regular classes would be necessary for class size reductions to be a cost-effective way of increasing student learning. This is a very large effect, and may be quite difficult to achieve.

TABLE 1

Correlation Between Individual Class Sizes and Stanford Achievement Test Scores in Reading, Math and Listening by Grade

		r	R^2	b	Deviation from Linearity
Kindergarten	Total Reading	-.16	.026	-.77*	NS**
	Math	-.12	.014	-.85	NS
First Grade	Total Reading	-.19	.037	-1.64*	NS
	Math	-.18	.031	-1.26*	NS
	Listening	-.17	.028	-.86*	NS
Second Grade	Reading	-.23	.052	-1.36*	NS
	Math	-.21	.044	-1.28*	NS
	Listening	-.22	.047	-1.05*	NS

Achievement scores are all scale scores in the Stanford Achievement Test series. Correlations are between class mean scores and class size. There was no listening test in the kindergarten series. b indicates the average decline in test score points for each pupil added to the class.

*Significant at $p<.01$.

**NS means the deviation from linearity was not significant at .10 level.

TABLE 2

Stanford Achievement Total Reading Scale Scores
By School Type, By Class Type, By Socioeconomic Status
Second Grade

Inner City Schools

	Low SES	n	High SES	n	Difference Low		Difference High
Small	564.9	(319)	587.7	(50)	15.8*	H >	28.4*
Regular	549.1	(449)	559.3	(36)	---		---
Reg w/Aide	552.6	(446)	580.7	(42)	3.5*	H >	21.4*

Suburban Schools

	Low SES	n	High SES	n	Difference Low		Difference High
Small	571.6	(158)	597.2	(276)	7.6*	< L	5.1*
Regular	564.0	(142)	592.1	(299)	--		---
Reg w/Aide	553.6	(174)	595.6	(320)	-10.4*	H >	3.5*

Rural Schools

	Low SES	n	High SES	n	Difference Low		Difference High
Small	582.7	(317)	608.9	(502)	6.8*	H >	9.4*
Regular	575.9	(403)	599.5	(576)	---		---
Reg w/Aide	582.8	(400)	602.5	(609)	6.9*	< L	3.0*

Urban Schools

	Low SES	n	High SES	n	Difference Low		Difference High
Small	579.7	(69)	609.4	(109)	10.9*	0	10.6*
Regular	568.8	(82)	598.8	(96)	---		---
Reg w/Aide	574.0	(92)	599.9	(129)	5.2*	< L	1.1

Numbers in parentheses indicate the number of students in each category. Among Low SES students the score differences favoring Small Classes over Regular Classes range from 15.8 points in Inner City Schools down to 6.9 points in Rural Schools. Among High SES students, the Small Class score advantage ranged from 28.4 to 5.1. The score advantage was higher for High SES students than for Low SES students in both Rural and Inner City Schools, the score advantage was higher for Low SES students in Suburban Schools, and there was no difference in Urban schools.

*Difference is statistically significant $p<.01$.

TABLE 3

Stanford Achievement Total Math Scale Scores
By School Type, By Class Type, By Socioeconomic Status
Second Grade

Inner City

	Low SES	High SES	Difference Low		Difference High
Small	564.0 (317)	587.7 (49)	9.8*	H >	19.6*
Regular	554.2 (449)	568.1 (36)	---		---
Reg w/Aide	553.7 (445)	582.4 (42)	0.5	H >	14.3*

Suburban

	Low SES	High SES	Difference Low		Difference High
Small	571.6 (157)	593.9 (276)	14.7	< L	8.7*
Regular	556.9 (142)	585.2 (300)			
Reg w/Aide	547.6 (172)	583.5 (318)	-9.3*	H >	1.7

Rural

	Low SES	High SES	Difference Low		Difference High
Small	580.7 (316)	600.9 (502)	3.2*	H >	4.4*
Regular	577.5 (401)	596.5 (575)	---		---
Reg w/Aide	583.7 (400)	596.7 (610)	6.2*	< L	0.2

Urban

	Low SES	High SES	Difference Low		Difference High
Small	566.8 (68)	605.6 (108)	11.3*	H >	19.9*
Regular	555.5 (82)	585.7 (96)	---		---
Reg w/Aide	565.9 (92)	589.8 (129)	10.4*	< L	4.1*

Numbers in parentheses indicate the number of students in each category. Three of the four comparisons of differences between Small and Regular students favor the High SES category. Two of four comparisons of Reg w/Aide and Regular favor High SES.

*Difference is statistically significant $p < .01$.

TABLE 4

End of Kindergarten to End of 2nd Grade
Stanford Achievement Scale Score Gains
by School Type, Class Type, and Socioeconomic Status
For Total Reading and Math

Total Reading

	Low SES						High SES					
	Small		Reg		R/A		Small		Reg		R/A	
Inner City	135	>	125	>	124		148	>	137	<	144	
Suburban	135	<	138	>	126		147	0	146	<	153	
Rural	151	>	148	<	156		164	>	159	0	160	
Urban	147	>	142	<	148		151	>	149	0	150	

Math

	Low SES						High SES					
	Small		Reg		R/A		Small		Reg		R/A	
Inner City	89	>	86	0	86		98	>	92	<	93	
Suburban	89	>	72	<	105		94	>	86	<	100	
Rural	100	<	104	<	110		101	<	106	>	104	
Urban	91	>	87	<	92		92	>	88	<	99	

Gain scores are computed by subtracting the average score in each category in kindergarten from the average score in second grade in the same category.